

基于改进密度聚类法的高压加热器传热系数研究

钱虹^{1,2}, 王海心¹, 张栋良^{1,2}

(1. 上海电力大学自动化工程学院, 上海 200090; 2. 上海市电站自动化技术重点实验室, 上海 200072)

摘要: 为了及时发现并处理高压加热器运行经济性失常, 采用传热系数直观地反映高压加热器的运行效率, 提出基于时序数据分析方法得到传热系数的在线动态模型。首先通过热动力学机理分析得到影响高压加热器传热系数的主要特征参数并建立基于特征参数的动态模型; 其次, 通过蜻蜓算法改进的密度聚类方法构建具有最优邻域参数的优化聚类模型, 得到可信端差区间。通过一段时间的某电厂的计算结果比较表明, 基于改进密度聚类法的传热系数在线动态模型在计算高压加热器传热系数时均方误差 MSE 低至 0.030 5%, 说明该模型有效、可行。

关键词: 高压加热器; 传热系数; 数据处理; 蜻蜓算法; 密度聚类

中图分类号: TH86 文献标识码: A DOI: 10.16146/j.cnki.rndlgc.2024.03.013

[引用本文格式] 钱虹, 王海心, 张栋良. 基于改进密度聚类法的高压加热器传热系数研究[J]. 热能动力工程, 2024, 39(3): 100-108. QIAN Hong, WANG Hai-xin, ZHANG Dong-liang. Study on heat transfer coefficient of high-pressure heater based on DA-DBSCAN[J]. Journal of Engineering for Thermal Energy and Power, 2024, 39(3): 100-108.

Study on Heat Transfer Coefficient of High-pressure Heater based on DA-DBSCAN

QIAN Hong^{1,2}, WANG Hai-xin¹, ZHANG Dong-liang^{1,2}

(1. College of Automation Engineering, Shanghai University of Electric Power, Shanghai, China, Post Code: 200090;
2. Shanghai Key Laboratory of Power Station Automation Technology, Shanghai, China, Post Code: 200072)

Abstract: In order to detect and deal with the operating economic abnormality of high pressure heater in time, this study used heat transfer coefficient to directly reflect the operating efficiency of high pressure heater, and put forward an online dynamic model of heat transfer coefficient based on time series data analysis method. First, the main characteristic parameters affecting the heat transfer coefficient of the high-pressure heater were obtained through thermodynamic mechanical analysis, and a dynamic model based on the characteristic parameters was established; second, an optimal clustering model with optimal neighborhood parameters was constructed by the density-based spatial clustering of applications with noise (DBSCAN) method improved by the dragonfly algorithm (DA) to obtain the credible end-difference interval. Through a period of comparative calculation results of a certain power plant, it is shown that the mean square error (MSE) of the heat transfer coefficient of the high-pressure heater is as low as 0.030 5% based on the online dynamic model of the heat transfer coefficient of the DA-DBSCAN, indicating that the model is effective and feasible.

Key words: high-pressure heater, heat transfer coefficient, data processing, dragonfly algorithm (DA), density-based spatial clustering of applications with noise (DBSCAN)

收稿日期: 2023-05-05; 修订日期: 2023-06-27

基金项目: 国家自然科学基金青年科学基金(51906133)

Fund-supported Project: Youth Fund of National Natural Science Foundation of China(51906133)

作者简介: 钱虹(1967-), 女, 上海电力大学教授。

引言

高压加热器是电站热力系统中的重要组成部分,其运行状况直接影响整个电站的热经济性。高压加热器的端差、压损、散热损失等因素对系统热经济性均有影响,其中端差的影响最大^[1-2]。研究高压加热器传热系数模型,对于及时发现并处理高压加热器运行经济性失常、提高核电站热经济性具有重要意义。本文利用加热器端差求解传热系数的计算方法,更客观、合理地监测高压加热器的热经济性。

高压加热器端差能够表征设备传热能效,在一定程度上可以反映设备运行热经济性波动趋势。闫顺林等人^[3]计算了高压加热器上端差的扰动对能效影响的强度系数。文献[4-5]分析端差对机组汽耗率、煤耗率的影响。谭良红等人^[6]提出了回热加热器变工况的机组热经济性通用模型,导出纯热量引入回热系统时的功率和热量变化通用计算公式。上述文章详细分析了端差对加热器经济性的影响,但未构建加热器端差的扰动对传热系数影响的数学模型。李勇等人^[7]结合热平衡方程通过传热系数计算低压加热器端差应达值,通过量化应达值与实际值的偏差,评价加热器性能。文献[8-9]在计算加热器端差应达值时认为,加热器传热系数不受汽轮机负荷、加热器入口水温等变化的影响,导致计算结果误差增大。进行传热系数计算需要给出正确描述设备状态的运行数据作为监测设备是否正常的标准,数据的处理非常关键。文献[10]提出一种基于密度的监测算法,该算法可以识别任意形状的数据集,但在处理密度分布不均的数据时聚类效果差。王晓龙等人^[11]构建了一种基于距离的异常检测方法,将超出安全运行域的数据判断为异常点,但需要手动切换工况,无法实时跟踪动态数据。文献[12]引入了局部离群因子(Local Outlier Factor, LOF),并通过数据的LOF值来判断该数据是否异常,此类算法只适用于静态的数据检测,不适合动态数据的检测过程。沈江飞等人^[13]结合神经网络模型对设备进行在线监测,解决了文献[12]存在的问题,但对于不同的参数需要重新调试,复杂度较

高。文献[14-15]基于聚类算法对数据集进行分类,在很大程度上能够检测出异常值,但均未考虑到复杂数据存在耦合关系,降低聚类效果。刘建东等人^[16]采用聚类算法对火电机组运行数据进行处理,划分为正常状态和泄露状态两类区域,实验结果表明,增加聚类算法有效提升监测准确度。文献[17]利用孤立森林算法(Isolation Forest)计算距离高密度数据群较远的数据点,计算孤立深度定义异常指标,对比异常指标完成异常判定。但此类算法未考虑数据点在序列中的先后顺序,难以识别序列中由于变化趋势导致异常的数据。然而,高压加热器的指标会随着设备的运行状态和劣化趋势发生改变,近期的数据相较于历史数据对计算当前高压加热器的传热系数更具有参考价值。因此,需要随着时间的推移确定各参数的正常运行区间,以确保传热系数计算的准确性。动态模型能够计算随序列均值改变而变化的时变阈值,更加注重序列近期数据的影响^[18]。

本文提出了一种基于改进密度聚类法(Dragonfly Algorithm-Density-Based Spatial Clustering of Applications with Noise, DA-DBSCAN)的高压加热器传热系数在线计算模型。首先,通过对高压加热器经济性分析,结合热动力学和热平衡方程,构建利用上端差和下端差求解传热系数的数学模型;其次,对高压加热器的时序数据进行处理,以滑动时间窗口的形式输入到DA-DBSCAN模型中,对上端差和下端差进行可行范围数据处理与均值计算,进而求得高压加热器的时变传热系数。实验结果表明,本文所提出模型简单有效,具有较高的聚类准确度,能够有效计算高压加热器传热系数。

1 基于热动力学的高压加热器传热系数模型

高压加热器是利用高压缸抽汽与给水进行热交换以提升给水温度的设备。当高压加热器运行异常时,会导致热交换不良,降低系统的经济性。目前,高压加热器多使用两段式高压加热器,其壳侧部分可以划分为疏水冷却段和蒸汽冷凝段^[19]。

高压加热器疏水冷却段传热系数表达式为:

$$K_1 = \frac{1}{\frac{1}{\alpha_{1d}} + \frac{\delta}{\lambda} + \frac{1}{\alpha_{2d}} \frac{d_o}{d_i} + R_d} \quad (1)$$

式中: K_1 —疏水冷却段的传热系数, $\text{kW}/(\text{m}^2 \cdot ^\circ\text{C})$; α_{1d} —疏水对管壁的放热系数, $\text{kW}/(\text{m}^2 \cdot ^\circ\text{C})$; δ —换热管壁厚, m ; λ —管壁传热系数, $\text{kW}/(\text{m}^2 \cdot ^\circ\text{C})$; α_{2d} —管壁对凝结水的放热系数, $\text{kW}/(\text{m}^2 \cdot ^\circ\text{C})$; d_o —换热管外径, m ; d_i —换热管内径, m ; R_d —疏水冷却段污垢热阻, $\text{m}^2 \cdot ^\circ\text{C}/\text{kW}$ 。

疏水冷却段热平衡方程:

$$D_{fw}(H'_{fw} - H_{fwi}) = D_c(H_{ec} - H'_{ec}) \quad (2)$$

$$K_1 A_{dc} \Delta t_{m1} = D_{fw} c_p (t'_{fw} - t_{fwi}) \quad (3)$$

式中: D_{fw} —加热器的凝结水流量, kg/s ; H'_{fw} —疏水冷却段出口焓值, kJ/kg ; H_{fwi} —给水入口焓值, kJ/kg ; D_c —抽汽流量, kg/s ; H'_{ec} —疏水冷却段疏水焓值, kJ/kg ; H_{ec} —蒸汽冷凝段疏水焓值, kJ/kg ; A_{dc} —疏水冷却段换热面积, m^2 ; Δt_{m1} —疏水冷却段的对数平均温差, $^\circ\text{C}$; c_p —高压加热器内蒸汽冷凝段凝结水的比定压热容, $\text{kJ}/(\text{kg} \cdot ^\circ\text{C})$; t'_{fw} —疏水冷却段出口水温, $^\circ\text{C}$; t_{fwi} —加热器的入口水温, $^\circ\text{C}$ 。

$$\Delta t_{m1} = [(t_{\text{sat}} - t'_{fw}) - (t_s - t_{fwi})] / \ln \frac{t_{\text{sat}} - t'_{fw}}{t_s - t_{fwi}} \quad (4)$$

式中: t_{sat} —高压加热器进口蒸汽压力下的饱和温度, $^\circ\text{C}$; t_s —高压加热器疏水温度, $^\circ\text{C}$ 。

高压加热器蒸汽冷凝段传热系数计算公式为:

$$K_2 = \frac{1}{\frac{1}{\alpha_{1n}} + \frac{\delta}{\lambda} + \frac{1}{\alpha_{2n}} \frac{d_o}{d_i} + R_n} \quad (5)$$

式中: K_2 —蒸汽冷凝段的传热系数, $\text{kW}/(\text{m}^2 \cdot ^\circ\text{C})$; α_{1n} —蒸汽冷凝段蒸汽对管壁的放热系数, $\text{kW}/(\text{m}^2 \cdot ^\circ\text{C})$; α_{2n} —管壁对凝结水的放热系数, $\text{kW}/(\text{m}^2 \cdot ^\circ\text{C})$; R_n —蒸汽冷凝段污垢热阻, $\text{m}^2 \cdot ^\circ\text{C}/\text{kW}$ 。

蒸汽冷凝段热平衡方程:

$$D_{fw}(H'_{fw} - H_{fwo}) = D_c(H_c - H_{ec}) \quad (6)$$

$$K_2 A_c \Delta t_{m2} = D_{fw} c_p (t_{fwo} - t'_{fw}) \quad (7)$$

式中: H_{fwo} —加热器出口焓值, kJ/kg ; H_c —抽汽焓值, kJ/kg ; A_c —蒸汽冷凝段换热面积, m^2 ; Δt_{m2} —蒸汽冷凝段的对数平均温差, $^\circ\text{C}$; t_{fwo} —加热器的出口水温, $^\circ\text{C}$ 。

$$\Delta t_{m2} = (t_{fwo} - t'_{fw}) / \ln \frac{t_{\text{sat}} - t'_{fw}}{t_{\text{sat}} - t_{fwo}} \quad (8)$$

加热器总体传热量方程:

$$KA \Delta t_m = k_1 A_{dc} \Delta t_{m1} + k_2 A_c \Delta t_{m2} \quad (9)$$

式中: K —加热器的总体传热系数, $\text{kW}/(\text{m}^2 \cdot ^\circ\text{C})$; A —加热器换热面积, m^2 ; Δt_m —高压加热器整体的对数平均温差, $^\circ\text{C}$ 。

$$\Delta t_m = [(t_{\text{sat}} - t_{fwo}) - (t_s - t_{fwi})] / \ln \frac{t_{\text{sat}} - t_{fwo}}{t_s - t_{fwi}} \quad (10)$$

将式(3)和式(7)代入至式(9)中,得到基于加热器进出口水温的传热系数 K 的数学模型:

$$K = \frac{D_{fw} c_p (t_{fwo} - t_{fwi})}{A \Delta t_m} \quad (11)$$

上端差、下端差和水位为表征高压加热器效率的指标^[20]。上端差 Δt_1 是指加热器进口蒸汽压力下的饱和温度 t_{sat} 与加热器给水出口温度 t_{fwo} 的差值, 下端差 Δt_2 是指加热器疏水温度 t_s 与加热器进口水温 t_{fwi} 的差值。上端差增大说明高压加热器热交换未达到设计要求。下端差无论增大还是减小, 当偏离设计值时会排挤下一级抽汽, 增加本级加热抽汽, 使得高压加热器的传热系数降低。由于水位高低会影响端差的大小, 水位的波动包含于端差的变化中, 故本文主要利用端差来计算传热系数。高压加热器端差对应的数学模型为:

$$\Delta t_1 = t_{\text{sat}} - t_{fwo} \quad (12)$$

$$\Delta t_2 = t_s - t_{fwi} \quad (13)$$

式中: Δt_1 —高压指加热器的上端差, $^\circ\text{C}$; Δt_2 —高压加热器下端差, $^\circ\text{C}$ 。

根据式(8)和式(12)可以推导出上端差会影响蒸汽冷凝段的对数平均温差; 根据式(4)和式(13)可以推导出下端差会影响疏水冷却段的对数平均温差。

通过上述分析构建基于高压加热器上、下端差的传热系数 K' 的数学模型为:

$$K' = \frac{K_1 A_{dc} (t_{\text{sat}} - t'_{fw}) - \Delta t_2}{A \Delta t_m \ln(t_{\text{sat}} - t'_{fw}) / \Delta t_2} + \frac{K_2 A_c (t_{fwo} - t'_{fw})}{A \Delta t_m \ln(t_{\text{sat}} - t'_{fw}) / \Delta t_1} \quad (14)$$

根据式(14)可知, 上端差、下端差可反映高压加热器的传热系数变化, 因此对上、下端差两个经济性指标进行数据处理, 能够有效提高传热系数计算的准确性。由于上、下端差标为机理复杂的二次参数, 在数据处理过程中会出现噪声或其他难以预见

的问题,故本文采用数据驱动方法对获得的上、下端差数据进行处理,剔除异常值,降低传热系数计算误差。

2 基于改进密度聚类法的时序参数优化处理

结合滑动时间窗口输入高压加热器上、下端差时序数据,采用 DA-DBSCAN 算法检测数据奇异值获取时序参数的可行范围,通过计算范围内的数据均值得到高压加热器传热系数。利用蜻蜓算法求解密度聚类方法的最佳邻域参数,提高奇异值判定的正确性,优化传热系数在线计算模型。

2.1 滑动时间窗口

由于核电系统采集到的数据是实时、动态,而传统的异常数据检测方法多应用静态数据,难以在动态的数据检测中发挥有效作用^[21]。因此,本文引入滑动时间窗口(Sliding Window, SW),将动态的数据流静态化处理,一旦时间窗口的宽度确定,即当前窗口内的数据也被确定。本文采集高压加热器两个经济性特征参数 Δt_i 的时序运行数据 $\Delta t_{i0} = \{(x_{i0}, t_0), (x_{i1}, t_1), \dots, (x_{in}, t_n)\}$, 其中 x_{ik} 为特征参数 Δt_i 在 t_k 时刻到达的数据,用 n 个滑动时间窗口 $(t_{i0}, t_{i1}, \dots, t_{ij}, \dots, t_{i(n-1)})$ 表示模型中使用到的所有样本集合 Δt_{i0} , 其中 t_{ij} 为特征参数 Δt_i 的第 j 个时间窗口,滑动时间窗口的长度设为 l 。在模型在线更新过程中滑动时间窗口随着时间推移向前移动。随着滑动时间窗口的移动样本空间的更新情况如图 1 所示。

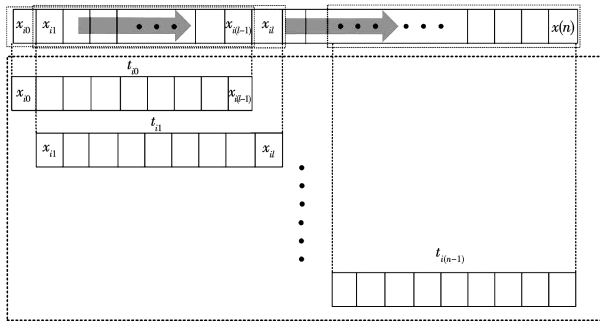


图 1 滑动时间窗口

Fig. 1 Sliding time window

随着时间序列改变得到对应的时变阈值和时变均值,是以当前一段时间窗口内的采样点为数据集

合,增强了邻近时间段内的数据对当前均值的影响,削弱较远数据对均值获取的干扰,对于评估设备的传热性能更贴合,且能够根据设备当前的运行状态剔除异常值,确保传热系数计算的准确性。

2.2 改进密度聚类算法

本文利用高压加热器历史数据 $\Delta t_i = \{x_{i0}, x_{i1}, \dots, x_{ip}\}$ 结合蜻蜓算法寻找聚类算法中各经济性指标对应的最优邻域参数 $(\varepsilon'_i, \text{MinPts}'_i)$, 其中 ε'_i 是样本的邻域距离阈值,描述了邻域范围的大小, MinPts'_i 是样本在距离为 ε'_i 的邻域内样本对象个数的阈值,描述了样本邻域密度的大小。蜻蜓算法(Dragonfly Algorithm, DA)的思想源于蜻蜓在自然界中特有的两种行为:在静态群中,蜻蜓可以分成几个亚群,飞越不同的区域,这是全局搜索的主要目标;在动态蜂群中,大量的蜻蜓组成大蜂群,向一个方向迁徙,有利于局部的开发利用^[22]。研究表明,任何群体的行为都遵循 3 个基本原则:分离(避免个体与邻居的碰撞)、对齐(个体之间的速度匹配)和凝聚(移动到邻居的平均位置)。同时,任何蜂群的主要目标都是生存,所有的个体都有捕食猎物和躲避天敌^[23] 的行为。这些行为的数学模型如下:

分离:

$$S_i = - \sum_{j=1}^N X - X_j \quad (15)$$

式中: S_i —第 i 只蜻蜓的分离量; X —当前蜻蜓个体的位置; X_j —第 j 个相邻蜻蜓个体的位置; N —相邻蜻蜓的数量。

对齐:

$$A_i = \frac{\sum_{j=1}^N V_j}{N} \quad (16)$$

式中: A_i —第 i 只蜻蜓的对齐量; V_j —第 j 个相邻蜻蜓个体的速度。

凝聚:

$$C_i = \frac{\sum_{j=1}^N X_j}{N} - X \quad (17)$$

式中: C_i —第 i 只蜻蜓的凝聚量。

捕食猎物行为公式:

$$F_i = X^+ - X \quad (18)$$

式中: F_i —第 i 只蜻蜓在寻找猎物时移动的距离;

X^+ —食物来源的位置。

躲避天敌行为公式:

$$E_i = X^- - X \quad (19)$$

式中: E_i —第 i 只蜻蜓在躲避天敌时移动的距离; X^- —天敌的位置。

本文假设蜻蜓的行为是分离、对齐、凝聚、捕食猎物和躲避天敌 5 种校正模式的结合。变换人工蜻蜓的位置,模拟其在搜索过程中的动作。步进向量被定义为 ΔX ,并使用以下公式进行更新:

$$\Delta X_{t+1} = (sS_t + aA_t + cC_t + fF_t + eE_t) + w\Delta X_t \quad (20)$$

式中: s —分离权重; a —对齐权重; c —凝聚权重; f —食物因子; e —敌人因子; w —惯性权重; t —迭代计数器。

计算步长向量后,位置向量计算如下:

$$X_{t+1} = X_t + \Delta X_{t+1} \quad (21)$$

蜻蜓需要在没有相邻蜻蜓的情况下,以随机行走(Levy flight)的方式绕搜索空间飞行,防止蜻蜓个体原地不动。

蜻蜓的位置由以下公式更新:

$$X_{t+1} = X_t + \text{Levy}(d) \times X_t \quad (22)$$

式中: d —位置向量的维数。

Levy fligh 的计算方法如下:

$$\text{Levy}(x) = 0.01 \frac{r_1 \sigma}{|r_2|^{1/\beta}} \quad (23)$$

式中: r_1, r_2 — $[0, 1]$ 中的两个随机数; β —常数。

$$\sigma = \left(\frac{\Gamma(1 + \beta) \times \sin(\frac{\pi\beta}{2})}{\Gamma(\frac{1 + \beta}{2}) \times \beta \times 2^{(\frac{\beta-1}{2})}} \right)^{\frac{1}{\beta}} \quad (24)$$

将蜻蜓算法运用到密度聚类算法邻域参数优化过程中,蜻蜓个体代表参数,参数个数代表个体维度,蜻蜓位置的移动代表参数数值的变化。蜻蜓与食物之间的距离用适应度函数值表示,函数值越小参数值越优。蜻蜓算法中的寻优函数没有固定形式,寻优过程是一个缩短个体与食物之间距离的过程,所以构造或选取的寻优函数必须具有极小值。采用 DB 指数作为适应度函数,建立模型如下:

$$\begin{cases} \min P = \frac{1}{N} \sum_{i,j=1}^N \max \frac{d_i + d_j}{d_{ij}} \\ 2\varepsilon \leq d_i, d_j, d_{ij} \leq \max d \end{cases} \quad (25)$$

式中: $\min P$ —个体与食物之间最小距离,即 DB 指数值; d_i, d_j —第 i, j 个类的直径。

结合约束条件对参数进行寻优,选择符合目标函数的参数组合作为最优邻域参数($\varepsilon'_i, \text{MinPts}'_i$),将得到的最优邻域参数输入到在线模型中。

密度算法(Density-Based Spatial Clustering of Applications with Noise, DBSCAN)将样本密度作为相似度,不断寻找被低密度区域分隔开的高密度区域,过滤噪声和离群点,以此划分样本簇群^[24]。密度聚类模型通过一对邻域参数($\varepsilon'_i, \text{MinPts}'_i$)对高压加热器样本的密度进行描述,采用欧氏距离计算方法描述两个数据点的距离 $D(x_i, x_j)$

$$D(x_i, x_j) = \sqrt{(x_j - x_i)^2} \quad (26)$$

式中: $D(x_i, x_j)$ —数据点 x_i 与 x_j 之间的距离。

通过将距离 $D(x_i, x_j)$ 与邻域参数 ε'_i 相比,判断是否属于同一簇群:

$$N_{\varepsilon'_i}(x_j) = \{x_i \in D \mid D(x_i, x_j) \leq \varepsilon'_i\} \quad (27)$$

同时,若数据点 x_j 的 ε'_i 邻域至少包含 MinPts'_i 个样本,即 $|N_{\varepsilon'_i}(x_j)| \geq \text{MinPts}'_i$,则代表 x_j 为一个核心对象,存在于 ε'_i 邻域内的数据点都是 x_j 的密度直达数据。密度聚类算法是根据密度可达原则去除样本数据的异常点,指从核心对象 x_{j1} 出发,途径一系列核心对象 x_{j2}, \dots, x_{jn} ,依次密度直达到 x_p ,那么 x_p 由 x_{j1} 密度可达。图 2 为密度聚类算法示意图。

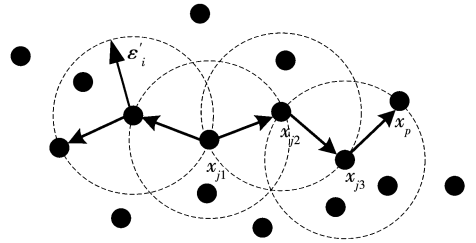


图 2 密度聚类算法示意图

Fig. 2 Schematic diagram of DBSCAN algorithm

本文采用 DB 指数评价聚类结果的准确性,DB 指数的核心思想是计算每个簇与最相似簇之间相似度,然后再通过求得所有相似度的平均值来衡量整个聚类结果的优劣^[25]。如果簇与簇之间的相似度越高(DB 指数偏高),也就说明簇与簇之间的距离越小,那么此时聚类结果就越差。

$$DB = \frac{1}{N} \sum_{i,j=1}^N \max R_{ij} \quad (28)$$

式中: $R_{ij} = \frac{s_i + s_j}{d_{ij}}$, s_i, s_j —第 i, j 个类的直径, R_{ij} 越小, 类之间的边界距离越远。 $DB \in [0, +\infty]$, 值越小表征分类效果越好。

对于高压加热器一组时间窗口内的数据样本 $t_{ij} = \{(x_{ij}, t_j), (x_{i(j+1)}, t_{j+1}), \dots, (x_{i(j+l)}, t_{j+l})\}$, 根据密度可达原则去除高压加热器数据样本的异常点, 得到滑动时间窗口数据样本 $d_{ij} = \{(x_{ij}, t_j), (x_{i(j+1)}, t_{j+1}), \dots, (x_{i(j+l)}, t_{j+l})\}$ 。再对样本 d_{ij} 进行处理, 求取高压加热器各经济性特征参数的均值为:

$$m_{ij} = \frac{x_{ij} + x_{i(j+1)} + \dots + x_{i(j+l)}}{l} \quad (29)$$

式中: l —数据样本总数。

动态下限:

$$A_{ij} = \max(x_{ij} + x_{i(j+1)} + \dots + x_{i(j+l)}) \quad (30)$$

动态上限:

$$B_{ij} = \min(x_{ij} + x_{i(j+1)} + \dots + x_{i(j+l)}) \quad (31)$$

若高压加热器指标实时数据超过正常运行区间 $[A_{ij}, B_{ij}]$, 表征当前值可能是一个坏值, 即不能正确反映设备当前运行状态, 且对监测设备传热性能会带来负面影响。应当在数据处理中剔除, 避免影响后续的传热系数计算有效性。

3 基于改进密度聚类法在核电站的应用实例

为验证基于改进密度聚类算法获取高压加热器传热系数模型的有效性和准确性, 将模型应用于7号高压加热器, 以获取端差均值并计算传热系数, 采用国内某电站1000 MW机组的500组现场运行数据作为实例数据集, 采样间隔为10 min。

3.1 蜻蜓算法寻找最优邻域参数结果

蜻蜓算法的参数设置: 种群数为30, 迭代次数为100, 邻域参数的范围依据经验设定, 在本文中给定上端差 Δt_1 的邻域参数范围为 $\varepsilon'_1 \in [0.01, 1]$, $\text{MinPts}'_1 \in [1, 5]$; 下端差 Δt_2 邻域参数范围为 $\varepsilon'_2 \in [0.01, 1]$, $\text{MinPts}'_2 \in [1, 5]$ 。基于蜻蜓算法寻找最优邻域参数的结果如表1所示。

表1 最优邻域参数

Tab. 1 Optimal neighborhood parameters

序号	ε'_1	MinPts'_1	ε'_2	MinPts'_2
1	0.030 589	1	0.036 808	1
2	0.039 666	1	0.037 072	1
3	0.035 729	1	0.043 724	1
4	0.027 577	1	0.039 381	1
5	0.034 281	1	0.043 276	1
6	0.034 879	1	0.043 134	1
7	0.030 907	1	0.037 109	1
8	0.041 749	1	0.038 581	1
9	0.032 051	1	0.037 031	1
10	0.035 540	1	0.035 454	1
Ave	0.034 297	1	0.039 157	1

3.2 模型对比

为了验证该模型的优越性, 本文将 DA-DBSCAN 聚类算法与基于遗传算法优化的密度聚类算法(GA-DBSCAN)、密度聚类算法(DBSCAN)、反向传播神经网络算法(BPNN)进行了对比。相关参数设置为: 在 DA-DBSCAN 算法中, 种群数为30, 迭代次数为100; GA-DBSCAN 算法的交叉概率为0.7, 突变概率为0.05, 种群数为30, 迭代次数为100; BPNN 算法中, 隐层节点数为15, 最大训练数为500, 显示间隔为10, 学习率为0.1。

经对比可知, 结合滑动事件窗口得到的动态阈值可以根据运行环境的改变做出调整, 直观表现出高压加热器经济性指标发生突变, 更准确地反映高压加热器传热性能变化。但不同的动态模型在数值处理中的性能也有差异。采用调整兰德系数(Adjusted Rand index, ARI)和 DB 指数(Davies-Bouldin Index, DB)作为模型的评价指标进行比较分析。表2为不同模型的聚类准确度对比。从表2可以看出, DA-DBSCAN 聚类结果的调整兰德系数高于 GA-DBSCAN、DBSCAN、BPNN 方法; DB 指数低于 GA-DBSCAN、DBSCAN、BPNN 方法。可见, 本文提出的模型聚类效果更好, 能更准确地获取正常数值。

表 2 不同模型的聚类准确度对比

Tab. 2 Comparison of clustering accuracy among different models

模 型	评价指标	计算次数					平均值
		1	2	3	4	5	
DA-DBSCAN	ARI	0.993 1	0.994 4	0.982 7	0.986 5	0.991 2	0.989 58
	DB	0.617 9	0.728 7	0.634 6	0.713 5	0.775 2	0.693 98
GA-DBSCAN	ARI	0.975 2	0.974 6	0.973 4	0.969 3	0.985 3	0.975 56
	DB	0.908 1	0.897 8	0.897 7	0.924 9	0.983 5	0.922 4
DBSCAN	ARI	0.912 7	0.903 0	0.908 4	0.905 7	0.904 6	0.906 88
	DB	2.463 5	2.348 4	2.846 2	2.947 3	2.856 2	2.692 32
BPNN	ARI	0.952 3	0.964 8	0.952 5	0.9614	0.938 5	0.953 9
	DB	1.394 7	1.372 5	1.237 4	1.304 5	1.384 6	1.338 74

采用本文方法对加热器进出口水温进行数据处理并结合式(11)可得到加热器实际传热系数 K ; 采用不同模型对上、下端差进行数据处理后得到的正常数值计算滑动时间窗口内的均值 m_{ij} , 结合式(14)绘制高压加热器传热系数 K' 曲线。选择均方误差 (MSE) 和均方根误差 (RMSE) 作为评估模型性能的指标。图 3 为传热系数曲线对比图。

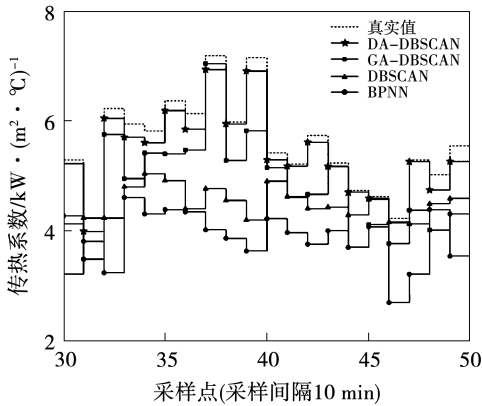


图 3 传热系数计算结果对比图

Fig. 3 Comparison graph of heat transfer coefficient calculation results

由图 3 中可见,使用本文所采用的 DA-DBSCAN 模型筛选出坏值后计算的传热系数与真实值更为接近,能快速跟随实际传热系数的变化趋势,表明该方法筛选坏值的效果更优异,模型可以用于准确计算高压加热器传热系数。从表 3 可见,DBSCAN 的效果相对最差,本文构建的 DA-DBSCAN 模型的计算误差最小,均方误差 MSE 低至 0.030 5%,均方根误差 RMSE 低至 0.174 6 $\text{kW}/(\text{m}^2 \cdot \text{°C})$,说明采用本文模

型利用端差计算高压加热器传热系数较为准确的。

表 3 不同模型的传热系数计算误差

Tab. 3 Calculation errors of heat transfer coefficients of different models

模型	MSE/%	RMSE/ $\text{kW} \cdot (\text{m}^2 \cdot \text{°C})^{-1}$
DA-DBSCAN	0.030 5	0.174 6
GA-DBSCAN	0.173 1	0.416 1
DBSCAN	0.573 0	0.757 0
BPNN	1.102 4	1.045 0

4 结 论

为监测高压加热器传热性能变化情况,本文提出了一种基于改进密度聚类算法获取高压加热器传热系数的在线动态模型,将该模型应用到实例中,得到如下结论:

(1) 基于加热器热平衡方程和传热数学模型,建立了高压加热器端差与设备传热系数的函数关系,说明通过端差计算高压加热器传热系数是行之有效的。

(2) 采用 DA-DBSCAN 来建立高压加热器传热系数在线动态模型,获取高压加热器端差可行范围内的时变均值并求解传热系数;同时针对密度聚类算法的邻域参数受主观影响的问题,以 DB 指数为评价指标结合蜻蜓算法进行寻优。经过实验,结果表明,本文构建的 DA-DBSCAN 模型所计算的高压加热器传热系数具有更高的准确性,均方误差 MSE 低至 0.0305%,均方根误差 RMSE 低至 0.104 8,均优于 GA-DBSCAN、DBSCAN 和 BPNN 方法。

参考文献:

- [1] 樊喜山. 热力发电厂加热器端差应达值初探[J]. 科技视界, 2013(20):161,179.
FAN Xi-shan. Initial investigation of end difference value obtained for thermal power plant heater[J]. Science & Technology Vision, 2013(20):161,179.
- [2] 张学镭,王松岭,陈海平,等. 加热器端差对机组热经济性影响的通用计算模型[J]. 中国电机工程学报, 2005, 25(4):168-173.
ZHANG Xue-lei, WANG Song-ling, CHEN Hai-ping, et al. General calculation model on the influence of terminal temperature difference on heat economy of unit[J]. Proceedings of the CSEE, 2005, 25(4):168-173.
- [3] 闫顺林,刘志巍,滕龙,等. 加热器上端差对机组能效影响的分析研究[J]. 汽轮机技术, 2014, 56(5):379-381,384.
YAN Shun-lin, LIU Zhi-wei, TENG Long, et al. Influence of heater's upper terminal temperature difference disturbances on energy efficiency[J]. Turbine Technology, 2014, 56(5):379-381,384.
- [4] 朱国栋,郑坚刚,李朋,等. 加热器端差对机组热经济性的影响[J]. 浙江电力, 2020, 39(6):74-80.
ZHU Guo-dong, ZHENG Jian-gang, LI Peng, et al. Influence of terminal temperature difference on thermal economy of generating units[J]. Zhejiang Electric Power, 2020, 39(6):74-80.
- [5] 王富兴,田松峰,麻东东,等. 加热器端差对机组煤耗影响简易计算[J]. 热力发电, 2011, 40(12):13-16.
WANG Fu-xing, TIAN Song-feng, MA Dong-dong, et al. Simple and easy calculation for influence of heater's terminal temperature difference upon coal consumption of 1 000 MW unit[J]. Thermal Power Generation, 2011, 40(12):13-16.
- [6] 谭良红,胡三高,曹晟磊. 基于回热加热器变工况的机组热经济性通用数学模型[J]. 华北电力大学学报(自然科学版), 2016, 43(4):104-110.
TAN Liang-hong, HU San-gao, CAO Sheng-lei. General mathematical model for heat economy of unit based on operation condition of regenerative heater[J]. Journal of North China Electric Power University, 2016, 43(4):104-110.
- [7] 李勇,马广辉. 一种评价低压给水加热器运行热经济性的方法[J]. 中国电机工程学报, 2012, 32(20):20-26,134.
LI Yong, MA Guang-hui. A method for evaluating thermal economic performance of low-pressure feedwater heaters[J]. Proceedings of the CSEE, 2012, 32(20):20-26,134.
- [8] 盛德仁,任浩仁,李蔚,等. 运行工况下汽轮机组主要参数应达值的数值分析[J]. 热力发电, 2000, 29(3):28-30,58.
SHENG De-ren, REN Hao-ren, LI Wei, et al. Primary parameter analysis of steam turbine unit under operation[J]. Thermal Power Generation, 2000, 29(3):28-30,58.
- [9] 任浩仁,李蔚,盛德仁,等. 火电机组变工况下运行指标应达值的分析[J]. 中国电机工程学报, 1999, 19(9):50-52,56.
REN Hao-ren, LI Wei, SHENG De-ren, et al. The analyze of operation index for the power unit under different loads[J]. Proceedings of the CSEE, 1999, 19(9):50-52,56.
- [10] 李少波,孟伟,璩晶磊. 基于密度的异常数据检测算法 GSWCLOF[J]. 计算机工程与应用, 2016, 52(19):7-11.
LI Shao-bo, MENG Wei, QU Jing-lei. GSWCLOF: Density-based outlier detection algorithm on data stream[J]. Computer Engineering and Applications, 2016, 52(19):7-11.
- [11] 王晓龙,张永发,刘忠,等. 基于数据驱动的核动力系统异常检测及分析方法研究[J]. 核动力工程, 2021, 42(5):149-155.
WANG Xiao-long, ZHANG Yong-fa, LIU Zhong, et al. Study on data driven anomaly detection and analysis algorithm for nuclear power systems[J]. Nuclear Power Engineering, 2021, 42(5):149-155.
- [12] 王茜,刘书志. 基于密度的局部离群数据挖掘方法的改进[J]. 计算机应用研究, 2014, 31(6):1693-1696,1701.
WANG Qian, LIU Shu-zhi. Improvement of local outliers mining based on density[J]. Application Research of Computers, 2014, 31(6):1693-1696,1701.
- [13] 沈江飞,李怀洲,黄立军,等. 基于局部离群因子和神经网络模型的设备状态在线监测方法研究[J]. 核动力工程, 2021, 42(3):160-165.
SHEN Jiang-fei, LI Huai-zhou, HUANG Li-jun, et al. Study on online monitoring of equipment condition based on local outlier factor and artificial neural networks model[J]. Nuclear Power Engineering, 2021, 42(3):160-165.
- [14] 周新宇. 基于耦合计算的聚类异常值检测方法研究[D]. 阜新:辽宁工程技术大学, 2022.
ZHOU Xin-yu. Research on clustering outlier detection method based on coupling calculation[D]. Fuxin: Liaoning Technical University, 2022.
- [15] 文雯,刘文哲,肖祥武,等. 基于大数据和并行随机森林算法火电机组供电煤耗计算模型[J]. 热力发电, 2018, 47(9):9-14.
WEN Wen, LIU Wen-zhe, XIAO Xiang-wu, et al. Calculation model of coal consumption for thermal power units based on big data and random forest algorithm[J]. Thermal Power Generation, 2018, 47(9):9-14.
- [16] 刘建东,丁启磊,张勇. 火电机组多能级系统泄漏智能在线监测系统的应用研究[J]. 热能动力工程, 2022, 37(7):27-33.
LIU Jian-dong, DENG Qi-lei, ZHANG Yong. Applied research on intelligent online leakage monitoring system for multi-level system of thermal power unit[J]. Journal of Engineering for Thermal Energy and Power, 2022, 37(7):27-33.
- [17] LIU F T, TING K M, ZHOU Z H. Isolation-based anomaly detection[J]. ACM Transactions on Knowledge Discovery from Data (TKDD), 2012, 6(1):1-39.
- [18] 王金明,王心培. 宏观经济极端景气变动的监测预警——基

- 于动态阈值模型[J]. 统计与决策, 2022, 38(20): 22-25.
- WANG Jin-ming, WANG Xin-pei. Monitoring and early warning of extreme prosperity fluctuation of macro economy — Based on dynamic threshold model [J]. Statistics & Decision, 2022, 38(20): 22-25.
- [19] 邹正宇. CANDU-6 核电厂系统与运行. 常规岛系统(一) [M]. 北京: 原子能出版社, 2010.
- ZOU Zheng-yu. CANDU-6 nuclear power plant system and operation. Conventional island system (Part I) [M]. Beijing: Atomic Energy Press, 2010.
- [20] 中国电力投资集团公司编著. 600 MW 火电机组节能对标指导手册[M]. 北京: 中国电力出版社, 2008.
- China Power Investment Corporation. Guidebook for energy efficiency benchmarking of 600 MW thermal power units [M]. Beijing: China Electric Power Press, 2008.
- [21] 张伊扬, 钱育蓉, 陶文彬, 等. 基于深度学习的属性图异常检测综述[J]. 计算机工程与应用, 2022, 58(19): 1-13.
- ZHANG Yi-yang, QIAN Yu-rong, TAO Wen-bin, et al. Survey of attribute graph anomaly detection based on deep learning[J]. Computer Engineering and Applications, 2022, 58(19): 1-13.
- [22] FENG Y, ZHANG P, YANG M, et al. Short term load forecasting of offshore oil field microgrids based on DA-SVM[J]. Energy Procedia, 2019, 158: 2448-2455.
- [23] 樊星男. 蜻蜓寻优算法在轴承故障模糊识别中的应用[J]. 机械设计与制造, 2023, 387(5): 31-36.
- FAN Xing-nan. Application of dragonfly optimization algorithm in fuzzy recognition of bearing fault[J]. Machinery Design & Manufacture, 2023, 387(5): 31-36.
- [24] 赫德军, 武欣嵘, 俞璐. 密度聚类方法研究[J]. 通信技术, 2022, 55(2): 135-142.
- HE De-jun, WU Xin-rong, YU Lu. Review of clustering algorithm based on density[J]. Communications Technology, 2022, 55(2): 135-142.
- [25] 王鹏龙. 一类高效的聚类有效性指标及应用[D]. 天津: 天津大学, 2014.
- WANG Peng-long. A kind of efficient clustering validity index and its application[D]. Tianjin: Tianjin University, 2014.

(姜雪梅 编辑)

(上接第 46 页)

- [8] 冯引利, 吴长波, 高鹏, 等. 某粉末冶金高温合金涡轮盘破裂转速分析[J]. 航空动力学报, 2013, 28(3): 501-506.
- FENG Yin-li, WU Chang-bo, GAO Peng, et al. Analysis of power metallurgy superalloy turbine disc's burst speed [J]. Journal of Aerospace Power, 2013, 28(3): 501-506.
- [9] 秦仕勇, 冯引利, 孙海鹤, 等. 粉末冶金涡轮盘破裂转速分析与验证[J]. 燃气涡轮试验与研究, 2017, 30(5): 36-41.
- QIN Shi-yong, FENG Yin-li, SUN Hai-he, et al. Burst speed analysis and verification of the powder metallurgy superalloy turbine disk [J]. Gas Turbine Experiment and Research, 2017, 30(5): 36-41.
- [10] 张春成, 汪好, 陈国光. 温度场对航空发动机转子超转破裂的影响[J]. 航空学报, 2019, 40(10): 131-138.
- ZHANG Chun-cheng, WANG Hao, CHEN Guo-guang. Impacts of temperature fields on rotor over-speed fracture of aircraft engines [J]. Acta Aeronautica et Astronautica Sinica, 2019, 40(10): 131-138.
- [11] 聂卫健, 邓旺群, 皮滋滋, 等. 民用涡轴发动机动力涡轮盘破裂转速研究[J]. 燃气涡轮试验与研究, 2021, 34(5): 33-38.
- NIE Wei-jian, DENG Wang-qun, PI Zi-zi, et al. Study on rupture speed of power turbine disk for a civil turboshaft engine [J]. Gas Turbine Experimental and Research, 2021, 34(5): 33-38.
- [12] 陈妍妍, 秦仕勇, 孙海鹤, 等. 修正平均应力法预测轮盘破裂转速的试验研究[J]. 航空动力学报, 2022, 37(5): 954-963.
- CHEN Yan-yan, QIN Shi-yong, SUN Hai-he, et al. Experiment study on predicting burst speed of disk by modified average stress method [J]. Journal of Aerospace Power, 2022, 37(5): 954-963.
- [13] 王开明, 王卫国, 陈国光. 基于极限应变法的超转分析验证与应用[J]. 推进技术, 2020, 41(7): 1594-1600.
- WANG Kai-ming, WANG Wei-guo, CHEN Guo-guang. Validation and application of over-speed analysis using strain-based fracture criteria [J]. Journal of Propulsion Technology, 2020, 41(7): 1594-1600.
- [14] 胡清清. 基于失效模式的轮盘破裂转速预测方法研究[D]. 杭州: 浙江大学, 2020.
- HU Qing-qing. Research on disk burst speed prediction method based on failure mode [D]. Hangzhou: Zhejiang University, 2020.
- [15] 《航空发动机设计手册总编委会》. 航空发动机设计手册第 18 册: 叶片轮盘及主轴强度分析[M]. 北京: 航空工业出版社, 2001.
- Chief Editing Organization of Aeroengine Design Handbook. Aeroengine design handbook (18th volume): Strength analysis of blade, disk and spindle [M]. Beijing: Aviation Industry Press, 2001.

(刘颖 编辑)